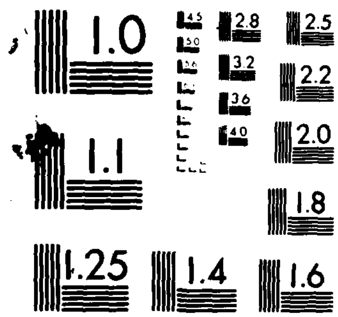
AD-A112 053 MASSACHUSETTS INST OF TECH LEXINGTON LINCOLN LAB F/G 17/2
A COMPARATIVE STUDY OF NARROWBAND VOCODER ALGORITHMS IN AIR FOR--ETC(U)
JAN 82 E SINGER F19628-80-C-0002
UNCLASSIFIED TR-590 ESD-TR-81-334 NL

END
DATE
FILMED
4 82
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

ESD-TR-81-[illegible]

Technical Report 590

E. Singer

A Comparative Study of Narrowband Vocoder Algorithms in Air Force Operational Environments Using the Diagnostic Rhyme Test

6 January 1982

Prepared for the Department of the Air Force
under Electronic Systems Division Contract F19628-80-C-[illegible] by

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Lexington, Massachusetts

The work reported in this document was performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology, with the support of the Department of the Air Force under Contract F19628-80-C-0002.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Raymond L. Loiselle

Raymond L. Loiselle, Lt.Col., USAF
Chief, ESD Lincoln Laboratory Project Office

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

A COMPARATIVE STUDY OF NARROWBAND VOCODER ALGORITHMS IN AIR FORCE OPERATIONAL ENVIRONMENTS USING THE DIAGNOSTIC RHYME TEST

E. SINGER
Group 24

TECHNICAL REPORT 590

6 JANUARY 1982

LEXINGTON MASSACHUSETTS

Abstract

This report presents a summary of work performed at Lincoln Laboratory aimed at improving the intelligibility of 2.4 kbps vocoders to be used in USAF operational environments. The distortions present in some of these environments, particularly the F-15 fighter aircraft, can place a severe burden on the speech modelling capabilities of contemporary vocoders. To study these effects and the benefits of various algorithmic improvements, the Diagnostic Rhyme Test was used as a means of providing an objective measure of relative system performance. A wide range of areas was explored through the use of real time computer simulations, including the effects of modified analysis and synthesis techniques, design parameter choices, interoperability, and environmental factors. The purpose of this report is to assemble and document the extensive body of DRT data which has been collected and thereby provide a means for the selection of design parameters likely to lead to improved vocoder performance.

CONTENTS

	Abstract	iii
I.	Introduction and Background	1
II.	Analysis	5
III.	DRT Scores	14
Acknowledgments		19
References		20

I. INTRODUCTION AND BACKGROUND

For the past two years, Lincoln Laboratory has been involved in a major effort to improve the quality and intelligibility of 2.4 kbps narrowband voice equipment to be used for Air Force air-to-air and air-to-ground communication. Although contemporary 2.4 kbps vocoders provide satisfactory performance when talkers are restricted to a relatively quiet, distortion-free environment, conditions in typical USAF airborne environments are considerably less benevolent. The combined influences of noise cancelling microphones, oxygen facemasks, aircraft audio systems, and high acoustic noise levels place a severe burden on the speech modelling capability of even the best narrow-band vocoders. A significant portion of the effort engaged in by Lincoln Laboratory has been directed toward an identification of the sources of degradation encountered within Air Force platforms and an evaluation of their effects on vocoder performance. As a result, extensive data has been gathered characterizing the noise field in the F-15 fighter aircraft in a variety of flight conditions[1], and the noisiest of these has been chosen as the basis for vocoder performance evaluation studies.

The measure chosen for quantification of system behavior is the Diagnostic Rhyme Test. Although no definite link between results of the DRT and user acceptability has yet been established, the DRT nevertheless provides a means for comparing the performance of a variety of systems in a repeatable and objective fashion. Three speakers, all former or active Air Force pilots, were chosen as subjects. Each speaker was required to read the DRT word lists while wearing the standard pilot headgear containing an oxygen facemask and

M101 noise cancelling microphone. The DRT data gathered using this talker base are self-consistent and provide a useful means for comparing the relative performance of the vocoder algorithms tested. However, other studies employing DRT results use a different talker base and considerable caution should be exercised when attempting to compare the absolute DRT scores contained in this report with those reported elsewhere.

The severity of the operating environment and the subsequent low DRT scores achieved using available narrowband algorithms led to an extensive and wide-ranging investigation of the many issues involved in vocoder design. Candidate algorithms were evaluated using signal processing digital computers which permitted the development of real-time vocoder simulations. It is the purpose of this report to assemble and document the extensive body of DRT data which has been collected and thereby provide a means for the selection of design parameters likely to lead to improved vocoder performance.

Preliminary experiments performed during the initial phase of the project indicated that the performance of modern LPC vocoders was severely compromised by the F-15 environment. Scores for these systems fell in the 70-75% DRT range. It was not clear whether the low score was the result of a deficiency in the linear prediction spectral modelling process or to a sub-optimum choice of design parameters. The need for a resolution of this issue led to the development of an experimental LPC vocoder incorporating a 90 Hz frame rate, 5 kHz audio bandwidth, and unquantized coefficients. This high quality vocoder achieved a DRT score of 84.7% and thus demonstrated that no fundamental deficiencies in the analysis-synthesis model existed which would preclude satisfactory vocoder operation in the heavily degraded F-15 audio environment. A similarly designed high quality channel vocoder produced about the same DRT score, indicating that a variety of analysis-synthesis methods are available for use under these conditions. Subsequently, considerable effort was directed toward a careful examination of the effect of the individual design parameters on vocoder performance and a determination of the relative contribution of each factor to intelligibility. These included signal conditioning, audio bandwidth, frame rate, LPC model order, and coding strategies. Many analysis techniques were evaluated, including time and frequency domain linear prediction, pitch-adaptive analysis, and high accuracy LPC parameter extraction. Synthesis methods which were explored involved continuous interpolation, multiple acoustic tubes, filter banks, and spectrum flattening. The DRT scores achieved by systems incorporating many of these variations are presented in Section III.

Another important element in this study was the restriction that modifications incorporated in any proposed 2.4 kbps vocoder not preclude interoperability with the proposed DoD narrowband system standard[2]. Implicit in this requirement is the fact that the data stream produced by a candidate system be consistent with that defined by the proposed DOD standard. An early experiment performed under this program concluded with the determination that the standard algorithm resident in the ITT Multi-Rate Processor terminal underwent a more severe loss in DRT (71.7%) than did the Lincoln Laboratory baseline non-interoperable system (75.2%). This result initiated a study aimed at providing an improvement capability within the limits of interoperability. A modest improvement was achieved by modifying the audio signal conditioning. Advanced synthesis techniques employing filter banks and spectrum flattening yielded a 2-3 point increase in the DRT score without affecting interoperability. A near-interoperable system which required replacement of the forward error control bits with parameter information produced another increase in the DRT. More sophisticated narrowband systems have also been demonstrated which, though non-interoperable, have achieved DRT scores above 80%.

Section II of this report represents an attempt to organize and highlight the results in a manner which illustrates some of the significant conclusions which have been drawn from the study. Of particular importance is the effect of model order and signal conditioning on interoperable vocoders. Also discussed are results relating to corrupting factors present in the F-15 environment such as the acoustic noise, oxygen facemask, and audio system. The results of experiments designed to determine the effects of audio bandwidth, frame rate, and coding on narrowband vocoders are considered next. Finally, the outcome of a series of experiments is presented in which various analysis and synthesis techniques are combined in an attempt to improve overall intelligibility. Section III presents a comprehensive listing of the DRT scores obtained thus far. It is hoped that the availability of this data base will prove useful to other investigators engaged in vocoder research.

II. ANALYSIS

Fig. 1 illustrates the results of a series of experiments designed to quantify the effects of various environmental factors on vocoder performance. The first set of scores (REF) relates to vocoder intelligibility under high quality noise-free conditions and serves as a reference against which other scores may be compared. The remaining scores demonstrate the influence on vocoder performance of the oxygen facemask, simulated F-15 acoustic noise, and a JTIDS Class 2 terminal audio card designed for an F-15 aircraft. The scores for the mask in noise-free conditions indicate that although unprocessed speech intelligibility suffers somewhat, the additional DRT loss due to narrowband analysis-synthesis is no different from that of the reference condition. Thus, the ability of LPC10 to model the speech signal does not appear to be compromised by the presence of the mask. However, the next set of data demonstrates the adverse effects of acoustic noise: although the decrease in intelligibility of the unprocessed speech is relatively small, the loss resulting from LPC10 is substantially increased. The final sets of scores illustrate that the presence of the low- and high-frequency rolloff characteristics introduced by the JTIDS audio card do not lead to a degradation in vocoder performance.

Fig. 2 attempts to illustrate the effects of two key features of the proposed DoD standard 2.4 kbps system on the intelligibility of processed speech in a simulated F-15 environment. A hardware implementation of the DoD standard as resident in the ITT Multi-Rate Processor (NSA LPC10 version 42) was evaluated along with a Lincoln Laboratory LPC10 vocoder

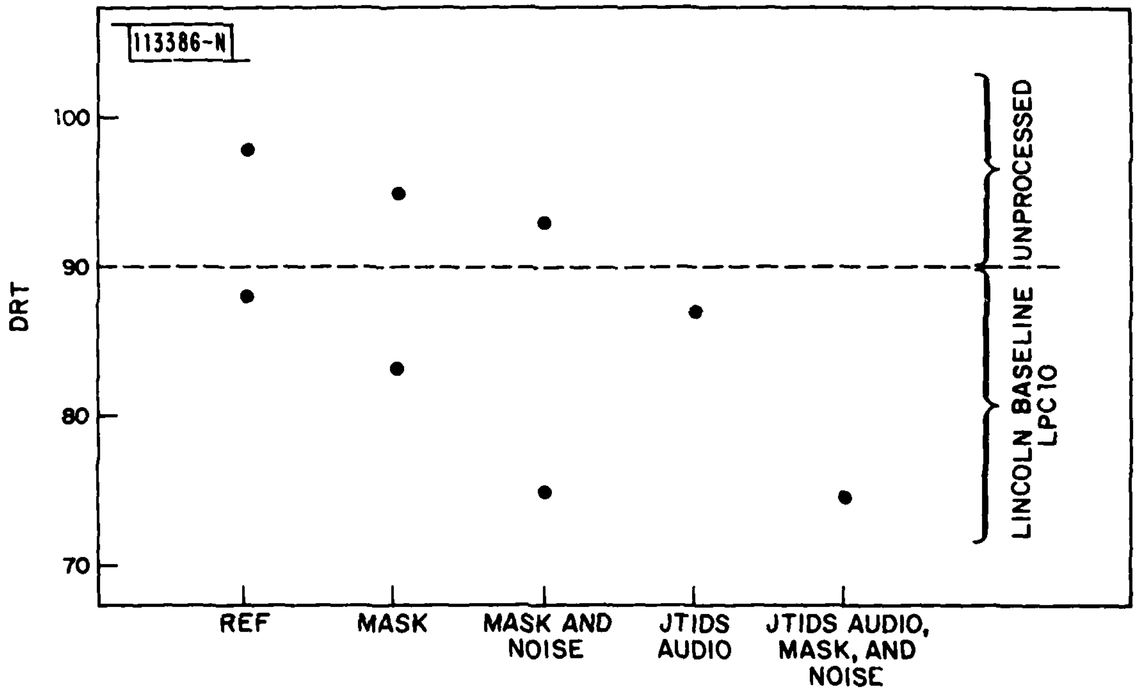


Fig. 1. Effects of environmental factors on vocoder intelligibility.

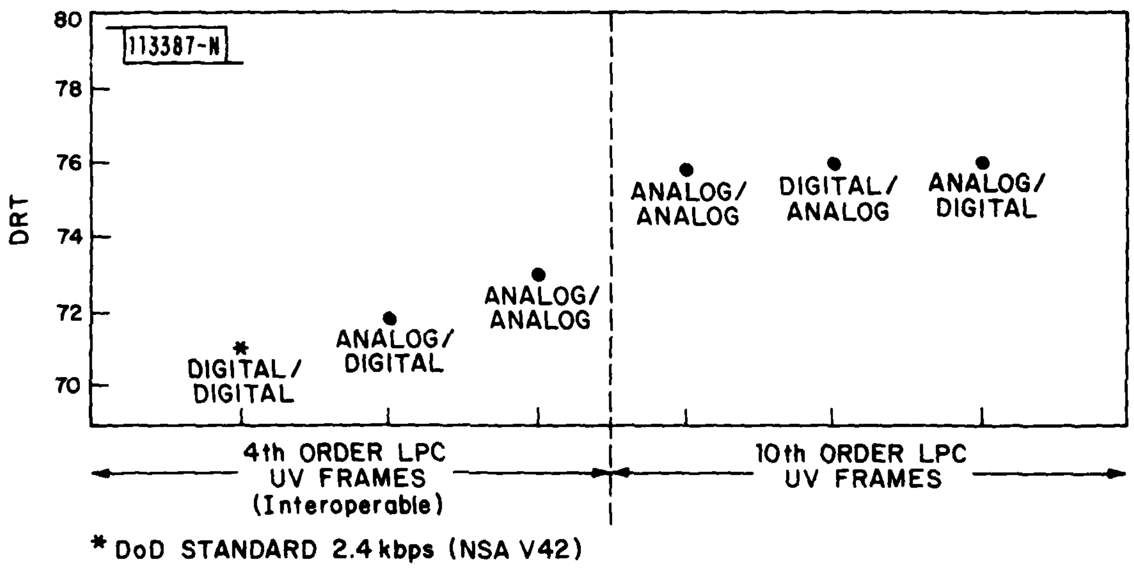


Fig. 2. Effects of signal conditioning and model order in F-15 environment.

implemented in software. The DoD standard performs digital pre-emphasis at the transmitter and digital de-emphasis at the receiver and applies a 4th order linear prediction spectral model to the speech signal during unvoiced frames. The DRT scores illustrate that the choice of signal conditioning in an interoperable system can lead to improved intelligibility, particularly if full analog is used. Furthermore, not only is a significant improvement in DRT scores obtained by using a full 10th order spectral fit during all frames, but the choice of signal conditioning within such a system is not critical. It is conjectured that although a reduced order model may be sufficient for use in conjunction with high quality input, the presence of high levels of ambient noise in the case of the F-15 requires higher order modelling capability. Also, the choice of model order depends on a voicing decision produced by the pitch detector under conditions where this system may not be totally reliable. It should be noted that while the 4th order systems are fully interoperable, those using a 10th order model technically are not.

Fig. 3 illustrates the effects of various design factors on a linear prediction analysis-synthesis system. These factors include frame rate, bit rate, audio bandwidth, model order, and parameter coding. The very last set of circles present the scores achieved by the Lincoln Laboratory baseline system and represent the starting point for this project. The leftmost scores were achieved by using increased frame rate and bandwidth (and hence model order). The bit rate of the high quality system was then reduced by applying the frame-fill strategy proposed by McLarnon[3]. The resulting 2.6 kbps vocoder produced a DRT score of 80.6% after coding. Since this system scores

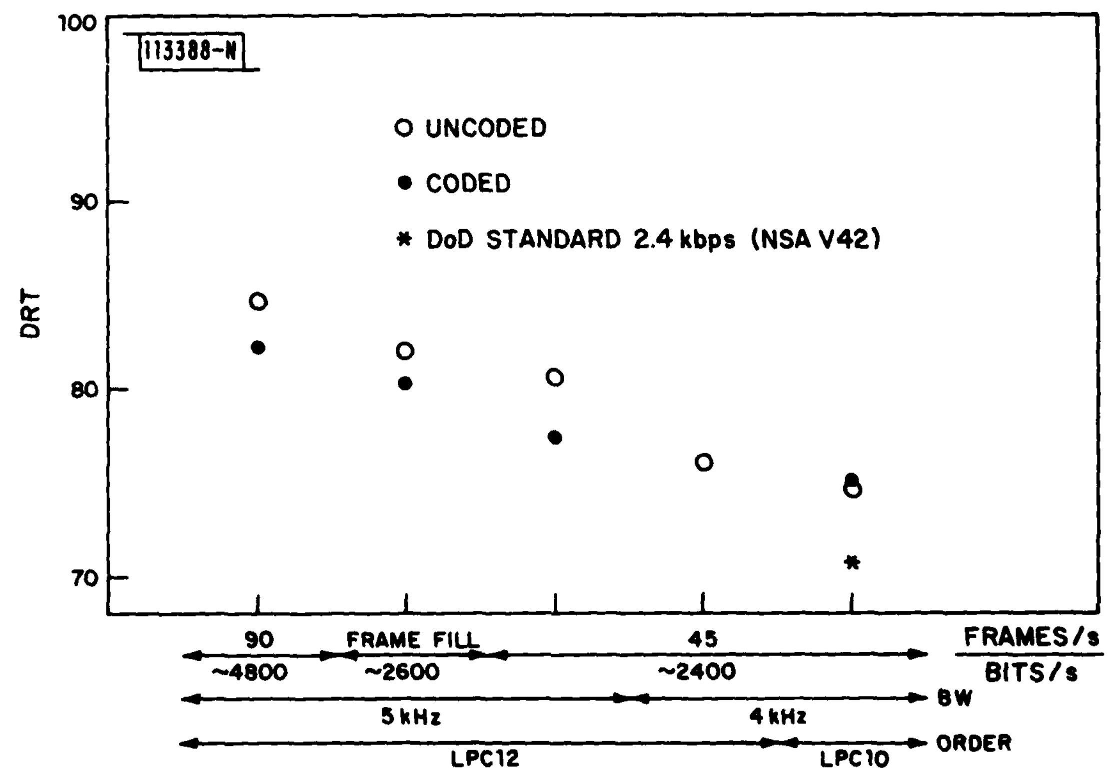


Fig. 3. Effects of frame rate, bandwidth, model order, coding, and frame fill in F-15 environment.

nearly 10 points better than the interoperable vocoder implemented in the ITT Multi-Rate Processor, it is considered to be the algorithm of choice for applications where interoperability is not critical. Work is currently in progress aimed at reducing the bit rate to 2.4 kbps using, for example, vector quantization schemes.

Fig. 4 can be used to summarize the performance of the "Extended Interoperable Systems", a term which has been adopted to describe a class of analysis-synthesis algorithms whose serial bit streams can be made to conform to the DoD standard but which do not utilize the conventional LPC analyzer or acoustic tube synthesizer. The first new analysis system to be tested was a frequency domain LPC algorithm[4] which, by using only the peaks of the high resolution spectrum, was expected to enhance the signal-to-noise ratio of the spectral measurements and hence improve the quality of the spectral fit. Fig. 4 presents a comparison of the DRT scores achieved by the frequency domain (FDLP/LPC) and standard time domain (LP/LPC) linear predictive analyses when combined with identical acoustic tube synthesizers. The results indicate that the frequency domain LPC technique does not produce an improvement in intelligibility.

In the next test, the frequency domain LPC analyzer (FDLP/LPC) was combined with a spectrally flattened channel vocoder synthesizer.[5] As shown in Fig. 4, the use of the flattened channel synthesizer introduced a three-point improvement in intelligibility as measured by the DRT. Subjective judgments obtained during informal listening indicated that the synthetic speech generated using the flattened channel synthesizer was of higher quality as well. Although this result was obtained in conjunction with

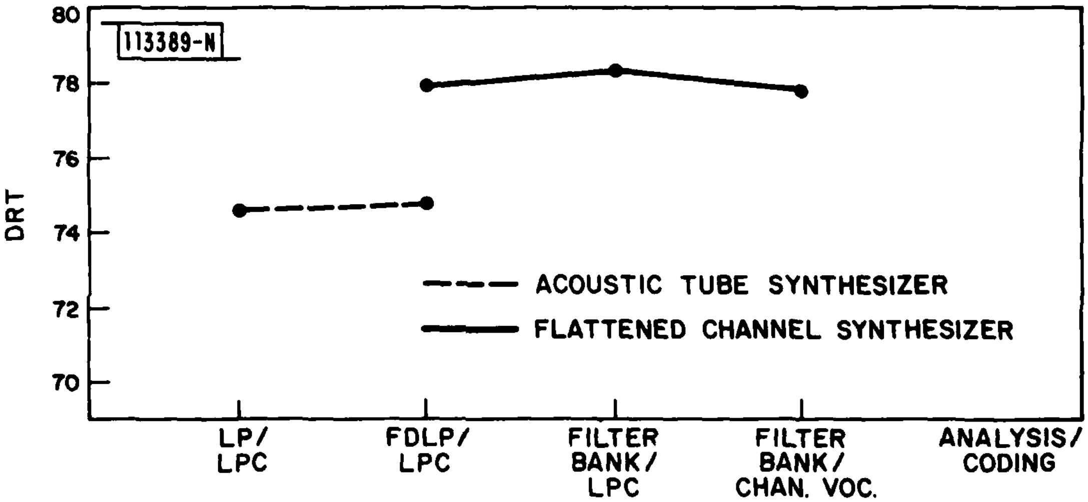


Fig. 4. Effects of analysis and synthesis techniques on vocoder performance in F-15 environment.

the frequency domain LPC analyzer, the evident correlation between the performance of the time and frequency domain methods of analysis suggests that the same improvement could be obtained using the time domain analyzer as well.

For the next experiment the flattened channel synthesizer was retained but a channel filter bank was used to provide the spectral information for the frequency domain LPC analyzer. The system (Filter Bank/LPC) produced a DRT score of 78.3, which, as shown in Fig. 4 is essentially the same as that achieved by the LPC analyzer (FDLP/LPC). In order to verify the fact that no information was lost as a result of using an all-pole interoperable spectral model to code the filter bank data, the results were compared with a high quality channel vocoder that uses the same 26 channel filter banks, the same flattened channel synthesizer, but standard channel spectrum coding at 4800 bps[6]. The score for this system (Filter Bank/Chan. Voc.) was essentially the same as those obtained using filter bank and frequency domain LPC methods. The results plotted in Fig. 4 demonstrate that while an improvement in DRT intelligibility can be achieved using advanced techniques at the synthesizer, both the linear prediction and filter bank analysis methods are equally effective in the F-15 environment. It is interesting to note that the use of the all-pole model to code the channel measurements reduces the channel vocoder data rate by a factor of 2.

As a result of the study of extended interoperable systems, the following conclusions may be drawn regarding narrowband vocoders operating in the F-15 environment:

1. An analyzer better than the standard time domain LPC algorithm has not yet been found.
2. The flattened channel synthesizer produces qualitatively and quantitatively better synthetic speech than the standard acoustic tube.

III. DRT Scores

Key to Terminology	
Term	*Description*
LPC10	Lincoln Lab Baseline LPC algorithm:
	- 10th order linear prediction using autocorrelation method
	- 180 sample non-overlapping analysis frame
	- Hamming window
	- 45 frames per second (fps)
	- Gold pitch detector
	- analog pre-emphasis and de-emphasis
	- acoustic tube synthesizer
	- non-interoperable
Interoperable	Lincoln Lab interoperable LPC10
Dynamic Microphone	GR 1960-9601 1/2" electret condenser microphone
Facemask	Air Force MBU-5/P oxygen facemask
M101	noise cancelling microphone used in oxygen facemask
Max. Likelihood Pitch	see Reference [7]
P&D	audio pre-emphasis and de-emphasis
Hitachi Audio	Hitachi HD44212 CODEC chip
AMI Audio	AMI S3505 CODEC chip
JTIDS Audio	SCI Systems JTIDS audio circuit card
Frame Fill	frame interpolation strategy [3]
LPCM	hardware vocoder [8]
Noise Prefilter	see Reference [9]
FDLPC	frequency domain linear prediction analysis [4]
ChanVoc	channel vocoder
Flattened ChanVoc	channel vocoder with spectrum flattening [5]
FlatVoc	high quality channel vocoder [6]
SEE	Spectrum Envelope Estimation [10]
DSVT	Digital Secure Voice Terminal
NSA V42	ITT Multi-Rate Processor using interoperable LPC10 (NSA Version 42)
F-15	F-15A high altitude, low level flight simulation [1]
Descent	F-15 gradual descent condition [1]
F-15A	second F-15A simulation [1]
F-16A	F-16A simulation [1]

Condition: QUIET, DYNAMIC MICROPHONE	*JH*	*PC*	*RM*	*Avg*
Baseline Systems				
Unprocessed	98.0 (0.54)	98.6 (0.44)	98.6 (0.27)	98.4 (0.41)
LPC10: Uncoded	90.1 (1.23)	91.1 (1.26)	86.1 (1.18)	89.1 (0.98)
LPC10	88.0 (1.15)	90.5 (1.25)	85.8 (1.16)	88.1 (0.79)
LPC12: Uncoded	85.0 (0.96)	93.1 (1.42)	88.5 (0.48)	88.9 (0.83)
Miscellaneous Systems				
LPC10: Max. Likelihood Pitch	89.1 (0.68)	92.4 (0.65)	84.5 (1.07)	88.7 (0.66)
LPC10: Pitch Adaptive Window	84.9 (1.56)	90.8 (1.63)	84.0 (1.54)	86.5 (1.49)
Interoperable	84.8 (1.54)	91.7 (0.98)	86.8 (1.29)	87.8 (0.97)
Audio Modifications				
LPC10: P&D out	87.5 (1.28)	89.1 (1.15)	83.2 (1.41)	86.6 (0.89)
LPC10: JTIDS audio card	87.2 (0.76)	87.8 (0.83)	86.8 (1.50)	87.3 (0.72)
LPC10: Hitachi audio	87.0 (0.48)	90.1 (1.02)	83.5 (0.91)	86.8 (0.59)

Condition: QUIET, FACEMASK	*JH*	*PC*	*RM*	*Avg*
Baseline Systems				
Unprocessed	93.4 (0.79)	94.5 (1.44)	96.2 (1.32)	94.7 (1.26)
LPC10: Uncoded	78.4 (1.74)	79.0 (0.95)	88.4 (1.07)	81.9 (1.16)
LPC10	79.0 (0.99)	81.2 (1.38)	88.4 (1.34)	82.9 (1.10)
LPC12: Uncoded	82.8 (1.04)	82.0 (1.00)	87.5 (0.94)	84.1 (0.83)
Miscellaneous Systems				
LPC10: Max. Likelihood Pitch	78.8 (1.66)	80.1 (1.52)	81.5 (2.12)	80.1 (1.56)
Audio Modifications				
LPC10: P&D out	77.7 (0.88)	80.5 (1.64)	75.8 (1.53)	78.0 (1.00)

Condition: QUIET, BOOM-MOUNTED M101	*JH*	*PC*	*RM*	*Avg*
Baseline Systems				
Unprocessed	94.4 (1.06)	97.0 (0.67)	94.9 (0.85)	95.4 (0.61)
LPC10	75.4 (1.61)	87.8 (1.02)	79.9 (1.51)	81.0 (0.92)

Condition: QUIET, FACEMASK+WINDSCREEN	*JH*	*PC*	*RM*	*Avg*
Baseline Systems				
Unprocessed	92.4 (0.67)		95.2 (0.79)	93.8 (0.74)
LPC10	75.1 (1.77)		78.0 (2.07)	76.6 (1.89)

Condition: F-15 (I)	JH	PC	RM	Avg
Baseline Systems				
Unprocessed	89.1 (1.02)	93.6 (0.82)	95.2 (0.62)	92.6 (0.69)
Unprocessed: 5kHz BW	86.7 (0.96)	91.7 (1.41)	91.8 (0.97)	90.1 (0.82)
Unprocessed: 3.8kHz BW	84.4 (0.56)	93.1 (0.62)	88.3 (1.16)	88.6 (0.53)
LPC10: Uncoded	69.1 (0.98)	78.9 (1.27)	75.8 (1.63)	74.6 (1.03)
LPC10	69.1 (1.35)	81.5 (1.82)	74.9 (1.25)	75.2 (1.19)
LPC12: Uncoded	71.1 (1.24)	83.1 (1.39)	74.5 (1.21)	76.2 (1.11)
LPC12: 5kHz BW, Uncoded	76.2 (1.03)	85.0 (0.98)	81.1 (1.01)	80.8 (0.66)
LPC12: 5kHz BW, 2.4kbps	70.4 (0.79)	83.6 (1.23)	79.4 (0.76)	77.8 (0.66)
LPC12: 5kHz BW, 90fps, Uncoded	79.2 (0.79)	86.8 (1.91)	88.0 (0.86)	84.7 (1.01)
LPC12: 5kHz BW, 90fps, 4.6kbps	77.7 (0.65)	86.6 (0.95)	82.6 (0.90)	82.3 (0.50)
Frame Fill				
LPC12: 5kHz BW, 90fps + Frame Fill, Uncoded	77.3 (1.69)	82.9 (0.96)	85.4 (0.81)	81.9 (0.73)
LPC12: 5kHz BW, 90fps + Frame Fill, 2.6kbps	74.9 (1.40)	83.3 (0.71)	83.5 (1.47)	80.6 (1.02)
Aliasing				
LPC10: 4kHz BW, 5kHz anl.filt., 4kHz syn.filt., 45fps, 2.4kbps	70.3 (1.87)	78.1 (1.59)	72.1 (1.32)	73.5 (1.41)
Audio Modifications				
LPC10: P&D out	72.4 (1.20)	81.2 (1.32)	73.8 (1.47)	75.8 (0.84)
LPC10: JTIDS audio card	69.8 (2.26)	80.2 (0.86)	72.3 (1.50)	74.1 (1.26)
LPC10: AMI audio	65.0 (1.52)	81.0 (2.08)	75.3 (1.60)	73.7 (1.33)
Noise Prefilter				
LPC10: Noise Prefilter	61.1 (2.07)	76.7 (1.45)	73.2 (1.51)	70.3 (1.30)
LPC10: Noise Prefilter, Max. Likelihood Pitch	63.2 (1.38)	77.7 (1.15)	74.9 (1.14)	71.9 (0.48)

Condition: F-15 (II)	*JH*	*PC*	*RM*	*Avg*
Interoperable Systems				
NSA V42	66.9 (1.47)	78.3 (0.85)	68.2 (1.42)	71.1 (1.01)
Interoperable	63.0 (1.74)	80.1 (2.82)	68.4 (1.92)	70.5 (1.95)
Interoperable: Analog P&D	67.1 (2.16)	80.1 (1.90)	71.7 (1.05)	73.0 (1.12)
Interoperable: Analog P, Digital D	64.8 (1.39)	74.7 (1.70)	75.9 (2.06)	71.8 (1.61)
Interoperable: Analog P, Digital D (alpha=0.75)	65.2 (2.17)	80.3 (1.86)	74.3 (2.51)	73.3 (1.84)
Interoperable: Analog P&D, 10 k's	68.5 (1.16)	82.6 (1.06)	75.9 (1.48)	75.7 (0.84)
Interoperable: Analog P, Digital D, 10 k's	70.8 (1.41)	81.9 (0.73)	75.4 (1.21)	76.0 (0.91)
Interoperable: Digital P, Analog D, 10 k's	70.4 (1.01)	82.6 (1.74)	74.9 (1.45)	76.0 (1.17)
Frequency Domain LP				
FDLPC10: 5kHz BW, Acoustic Tube syn., uncoded	70.7 (1.84)	80.3 (2.00)	78.0 (1.83)	76.3 (1.31)
FDLPC10: 5kHz BW, Flattened ChanVoc syn., uncoded	76.6 (1.55)	83.3 (1.38)	85.0 (1.45)	81.6 (0.99)
FDLPC10: 5kHz, Split Band anl., Flattened ChanVoc syn., uncoded	77.1 (1.65)	78.9 (1.50)	81.5 (2.08)	79.2 (1.47)
FDLPC10: 4kHz BW, Flattened ChanVoc syn., uncoded	72.8 (2.26)	78.9 (1.61)	81.9 (1.38)	77.9 (1.47)
ChanVoc: 4kHz BW, LP coef., uncoded	70.4 (1.68)	86.1 (0.92)	78.5 (0.52)	78.3 (0.79)
FDLPC10: 4kHz BW, Acoustic Tube syn., uncoded	71.7 (1.57)	77.2 (1.97)	75.7 (1.23)	74.9 (1.18)
High Quality Channel Vocoder				
FlatVoc: 5kHz BW, 100fps, 8kbps	80.1 (2.03)	87.4 (0.82)	83.3 (1.84)	83.6 (1.19)
FlatVoc: 5kHz BW, 100fps, Frame Fill	78.0 (1.60)	85.3 (1.33)	79.4 (0.83)	80.9 (1.15)
FlatVoc: 4kHz BW, 50fps, 4kbps	72.9 (1.43)	85.5 (1.12)	75.1 (1.64)	77.9 (0.91)
Telephonics M101				
Unprocessed: New M101		93.5 (1.18)	89.8 (1.11)	91.7 (1.03)
LPC10: New M101		83.2 (1.75)	72.5 (1.30)	77.9 (1.03)
Miscellaneous Systems				
LPCM	67.1 (1.13)	83.2 (1.19)	69.4 (2.35)	73.2 (1.19)
LPCM: P&D out	71.5 (1.40)	76.0 (1.42)	79.0 (1.37)	75.5 (1.18)
LPC10: Max. Likelihood Pitch	66.9 (1.18)	78.4 (1.30)	71.4 (0.96)	72.2 (0.82)
LPC10: Pitch Adaptive Window	65.8 (2.59)	80.5 (2.08)	79.8 (1.62)	75.2 (1.66)
SEE	66.8 (1.57)	79.2 (1.04)	71.6 (1.40)	72.5 (0.88)
CVSD16	72.1 (1.00)	83.3 (2.30)	83.2 (1.57)	79.6 (1.02)
CVSD16 (DSVT)	74.1 (1.30)	88.9 (1.02)	83.6 (1.39)	82.2 (0.94)

Condition: DESCENT	JH Live	JH Simulated
Baseline Systems		
Unprocessed	92.2 (0.86)	91.4 (1.36)
LPC10	74.2 (1.71)	77.1 (1.18)
Audio Modifications		
LPC10: P&D out	74.6 (1.35)	75.4 (1.20)
Miscellaneous Systems		
LPCM	75.9 (1.84)	73.8 (1.27)
LPCM: P&D out	78.3 (1.30)	75.0 (1.60)
NSA V42	77.5 (1.74)	
FlatVoc: 5kHz BW, 8kbps	83.2 (1.68)	

Condition: DESCENT+WINDSCREEN	JH Simulated
Baseline Systems	
Unprocessed	93.4 (0.48)
LPC10	76.2 (1.07)
Audio Modifications	
LPC10: P&D out	72.7 (0.98)

Condition: F-15A	JH	PC	RM	Avg
Baseline Systems				
Unprocessed	81.0 (1.68)	94.4 (0.86)	94.5 (1.06)	90.0 (1.05)
LPC10	57.9 (1.88)	82.3 (1.64)	76.2 (1.12)	72.1 (1.15)

Condition: F-16A	JH	PC	RM	Avg
Baseline Systems				
Unprocessed	85.3 (1.24)	95.7 (0.89)	96.6 (0.73)	92.5 (0.70)
LPC10	64.2 (1.61)	83.7 (0.79)	79.0 (1.34)	75.7 (0.90)

ACKNOWLEDGMENTS

I would like to express my gratitude to J. Tierney for his assistance in drafting this report and Dr. R. J. McAulay for his contributions to Section II. Producing the DRT data presented in Section III required the generation of a considerable number of recordings and a major portion of this effort was conducted by M. L. Malpass, Dr. R. J. McAulay, and Dr. B. Gold.

REFERENCES

1. R. L. Miller, R. D. Bruce, F. N. Iacovino, and A. W. F. Huggins, "Simulation of Cockpit Noise Environments in Four Tactical Aircraft for the Purpose of Testing Speech Intelligibility," Bolt, Beranek and Newman, Inc., Cambridge, MA, BBN Rpt. 4465 (September 1981).

2. Appendix I of ANDVT Procurement Specification, "LPC Processor Algorithm for the ANDVT Tactical Terminal (TACTERM)," Specification No. TT-B1-4210-0087 (May 1978).

3. E. McLarnon, "A Method for Reducing the Transmission Rate of a Channel Vocoder by Using Frame Interpolation," in Proc. Intl. Conf. Acoust., Speech and Signal Processing, Tulsa, Oklahoma (1978).

4. R. J. McAulay, "Maximum Likelihood Spectral Estimation of Peraodic Processes and Its Application to Narrowband Speech Coding," First ASSP Workshop on Spectral Estimation, McMaster University, Hamilton, Ontario, pp. 7.4.1-7.4.4 (August, 1981).

5. B. Gold and J. Tierney, "Pitch-Induced Spectral Distortion in Channel Vocoders," J. Acoust. Soc. Amer. 35, 730-731 (1963).

6. B. Gold, P. E. Blankenship, and R. J. McAulay, "New Applications of Channel Vocoders," IEEE Trans. Acoust., Speech and Signal Processing ASSP-29, 13-23 (1981).

7. R. J. McAulay, "Design of a Robust Maximum Likelihood Pitch Estimator for Speech in Additive Noise," Technical Note 1979-28, Lincoln Laboratory, M.I.T. (June 1979), DDC-AD-A077159.

8. E. M. Hofstetter, J. Tierney, and O. Wheeler, "Microprocessor Realization of a Linear Productive Vocoder," IEEE Trans. Acoust., Speech and Signal Processing ASSP-25, 379-387 (1977).

9. R. J. McAulay and M. L. Malpass, "Speech Enhancement Using a Soft-Decision Noise Suppression Filter," IEEE Trans. Acoust., Speech and Signal Processing ASSP-28, 137-145 (1980).

10. D. B. Paul, "The Spectral Envelope Estimation Vocoder," IEEE Trans. Acoust., Speech and Signal Processing ASSP-29, 786-794 (1981).

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
ESD-TR-81-334

2. GOVT ACCESSION NO.
AD-A112 053

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
A Comparative Study of Narrowband Vocoder Algorithms in Air Force Operational Environments Using the Diagnostic Rhyme Test

5. TYPE OF REPORT & PERIOD COVERED
Technical Report

6. PERFORMING ORG. REPORT NUMBER
Technical Report 590

7. AUTHOR(s)
Elliot Singer

8. CONTRACT OR GRANT NUMBER(s)
F19628-80-C-0002

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Lincoln Laboratory, M.I.T.
P.O. Box 73
Lexington, MA 02173

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
Program Element Nos. 27417F, 28010F and 33401F
Project Nos. 2283, 411L, 2264 and 7820

11. CONTROLLING OFFICE NAME AND ADDRESS
Air Force Systems Command, USAF
Andrews AFB
Washington, DC 20331

12. REPORT DATE
6 January 1982

13. NUMBER OF PAGES
28

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
Electronic Systems Division
Hanscom AFB, MA 01731

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

narrowband vocoder algorithms
fighter aircraft environment simulation
vocoder performance in simulated environments
vocoder design parameters
speech intelligibility

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report presents a summary of work performed at Lincoln Laboratory aimed at improving the intelligibility of 2.4 kbps vocoders to be used in USAF operational environments. The distortions present in some of these environments, particularly the F-15 fighter aircraft, can place a severe burden on the speech modelling capabilities of contemporary vocoders. To study these effects and the benefits of various algorithmic improvements, the *Diagnostic Rhyme Test* was used as a means of providing an objective measure of relative system performance. A wide range of areas was explored through the use of real time computer simulations, including the effects of modified analysis and synthesis techniques, design parameter choices, interoperability, and environmental factors. The purpose of this report is to assemble and document the extensive body of DRT data which has been collected and thereby provide a means for the selection of design parameters likely to lead to improved vocoder performance.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DAT
FILM